DOCUMENT RESUME

ED 252 571 TM 850 044

AUTHOR Gougeon, Deborah J.
TITLE The Limitations of SAT Math Scores in Predicting College Math Performance.
PUB DATE Nov 84
NOTE 14p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (13th, New Orleans, LA, November 14-16, 1984).
PUB TYPE Viewpoints (120) -- Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Admission Criteria; *College Entrance Examinations; Higher Education; High School Seniors; *Mathematics Achievement; Predictive Measurement; *Predictive Validity; Quantitative Tests; Testing Problems; Undergraduate Students
IDENTIFIERS *Scholastic Aptitude Test

ABSTRACT

This study was designed to determine if the decline evidenced in the College Entrance Examination Board (CEEB) Scholastic Aptitude Test (SAT) scores for mathematics is a true reflection of a decrease in students' mathematical ability, or is a result of the inadequacy of the test in measuring students' mathematical abilities. The data consisted of SAT mathematics test scores as well as the total mathematics grades of each student over the four-year college experience. The study suggested that SAT math scores are not a good predictor of college performance. The SAT mathematics test scores are often used erroneously as exclusive measurements of individual, or institutional, quality. A more valid procedure for understanding college achievement by the utilization of previous achievement measures for purposes of selection, placement, and guidance needs to be developed. There is also a need for methods to measure personality traits and abilities not measured by nationally normed tests. Other methods of improving the predictability factor of nationally normed tests are also recommended. (DWH)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED252571

Deborah J.Gougeon, Ph.D.
School of Management
University of Scranton
Scranton, Pa. 18510

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

SCOPE OF INTEREST NOTICE
The Eric Facility has assigned this document for proccesing to: TM

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should refect their special points of view SE

# The Limitations of SAT Math Scores in Predicting College Math Performance

In a recent UPI article entitled "College Board Frowns on SAT Comparisons," Robert Seaver, director of communication for the College Board, states, " [SAT] scores are intended primarily to help predict the academic performance of individual students during their first and subsequent college years." (Oct. 28) Indeed there seems to be a wide consenus in the educational community, and in society at large, that the SATs do exactly this. Consequently, the continuing decline in SAT scores nationally has been a cause of general concern. Not only are math scores decreasing from year to year but the situation appears to be getting worse rather than better (Oldenquist, 1983). Despite a small increase in last year's test results,( three points) average scores presently remain substantially lower than in the past.In the testing year 1962-1963 the average SAT math score was 502. Last year (1983-84 academic year) the average was 471 (Biemiller, 1984). Some educators believe that a decline in nationally-normed mathematics scores is an indication that our secondary educational system is either failing to teach mathematics courses properly or that other less difficult courses are being substituted for mathematics courses. Others feel that with a decline in the SAT math scores, students will be doing

TM 850 044

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
D.J. Gougeon
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

poorly in mathematics courses in U.S. colleges and universities thus furthering the trend toward mathematical illiteracy. In light of all this one might ask the questions: Are the CEEB SAT math scores really an indicator of future mathematics performance, and, ultimately, the future mathematics ability of our young people, as Mr. Seaver maintains, or, are the CEEB SAT mathematics scores a means which do not adequately test the ability of many students and therefore are not a good indicator of future college mathematics performance?

Interestingly enough, almost all of the validity data concerning the SAT scores have been reported by the Educational Testing Service itself (Whitla & Zimmerman, 1965). Despite the enormous power the ETS has over people's lives, and despite many questions that have arisen about the validity of multiple-choice exams in predicting college and career performance, and also about racial and economic biases in the tests, ETS has operated since 1947 totally free from government scrutiny or supervision. This situation has been a cause of concern for some. Others, for ideological reasons, dislike the near-monopoly of the testing business by the ETS and the secrecy with which this firm and others go about their business. Some dislike testing per se, viewing it more as a barrier to advancement than as an opportunity for objective evaluation. These concerns are indeed justified because of the enormous influence that such tests have.

The CEEB-SAT mathematics tests were designed to predict future college achievement among those students taking the tests. In addition to this screening device, some college mathematics departments have developed their own entrance tests. However, the

nationally-normed tests have long remained the major criteria for future college prediction. For more than half a century, this standarized test has been used as an aid in predicting future college achievement. A standarized test has been defined as one that implies uniformity of procedure in administering and scoring (Anastasi). In this sense, the CEEB-SAT mathematics test is standarized. However, it has been suggested that the test no longer possesses the predictive validity it once had since the performance on which the test norms are based may no longer be representative of the mathematical abilities and skills of the people now taking the tests. It may not be possible, therefore, to accurately compare test results from past & present test-taking populations. For example, according to one researcher (Wirtz), the recent increase in the number of women and low income students as candidates for college may have drastically changed the predictive capacities of these tests because test norms were established on the basis of the achievements of a particular group having characteristics much different from the characteristics of the present test-taking population. Because of this the use of nationally-normed tests for the prediction of college mathematics success could be working to the disadvantage of candidates for admission to colleges and universities,to the higher education institutions themselves,and ultimately,to all of society. If students are being admitted or rejected on the basis of predictors which may no longer be functioning properly, the interests of the educational establishment and society at large are not being properly served.

In addition to this limitation other researchers have cited

common misuses of such tests (Traxler). Among these are: 1) teacher quality and effectiveness are appraised by standarized tests without controlling extraneous variables, thus reducing confidence in tests; 2) there is an over emphasis on the tests themselves, sometimes resulting in students being coached for them; 3) much importance is sometimes attached to small differences in scores, even when the differences are not significant.

Despite these limitations, however, the current most widely accepted indicator of success or decline of mathematical achievement is still the CEEB-SAT mathematics test. This decline has in turn given rise to many incomplete, and sometimes questionable, speculations over the years as to what supposedly went wrong with American education. One study shows that items such as too much TV, old-fashioned (or new-fashioned) curricula, the decline of the family, Vietnam, the elimination of school prayer, and government interference have all been indicted (Oldenquist, 1983). The report from the Advisory Panel on the SAT Test Score decline states that the decline has developed in "2 distinct stages, characterized by significantly different balances of ... different causal factors." In the first stage, 1963 to about 1970, the report says, as much as three-quarters of the decline was caused by the "notable extension and expansion of educational opportunity in the U.S. during that period." This statement refers to sharp increases in the proportions of students taking the test who came from groups that have always registered substantially lower-than-average scores on the test.

As noted earlier, these are largely students from families with lower socioeconomic status, minorities, and women.

In the second stage of decline, after about 1970, there were fewer changes in the demographic distribution of the SAT takers. During this time, the decline in scores swept across the board: scores of traditionally higher-scoring students dropped, and those of traditionally lower-scoring students fell even lower. No single reason has emerged as the reason for the test score decline (Ryor, 1977). Wherever the blame is placed, the CEEB SAT math scores have remained as the primary means of measuring the declining mathematical ability of American students. Therefore, one would think that a moderate to high level of correlation must exist between CEEB SAT math scores and overall mathematics performance, if indeed a decline in mathematical ability has actually occurred. Studies in the area have shown mixed results up to this point.

A review of the literature shows that many studies have noted the relationship between SAT math scores and freshman mathematics grade point average while generally ignoring the remaining three years (Fricke,1965). Thus, while in one study SAT math scores yielded a substantial correlation of .62 with earned grades in freshman mathematics which were statistically significant beyond the .01 level (Gussett, 1974), another study in 1976 (Dalton) found that the ability to predict college grades from SAT math scores and high school average appeared to be decreasing. Some researchers have shown moderate to high correlations with SAT math scores and overall college performance

but in these cases other factors had been combined with the SAT as a predictor (Larson and Scontrino, 1976; Chissom and Lanier, 1975; Mauger and Kolmodin, 1975; Siegleman, 1977). The College Board contends that, based on its own research on the continuing validity of the test, 1) the ability of the test to forecast students' college grades has held up during the decline; 2) the test has not become harder during the decline; and 3) the decline does not result from more minorities and women taking the test.

Despite these assurances, and in response to a growing concern about the influence and validity of SAT tests, in 1979, New York State Senator Halperin co-sponsored a pioneering bill that would open standarized admissions tests to public scrutiny. Supporters claimed that it would expose test questions that are culturally biased and reveal errors in scoring. ETS officials, for the most part, dismissed the complaints and indicated bafflement at what the fuss was all about. They claimed with some justification that hard evidence of ETS's failings was sorely lacking. The firm is successful,they maintained because it was shrewd enough to fulfill the needs of disparate educational institutions at the right time; it has dominated the field because it hires the top experts and does its job well. It is somewhat secretive about its work because tests have to be secretive to be confidential. Despite this defense the bill passed and critics of the SAT continue to be vocal. Some criticisms have been directed against the format of the test itself. For example, a mathematician at Queens College in New York has written that mutiple-choice exams of the variety prepared by ETS "are inherently superficial, using vagueness,

imprecision, and worse as substitutes for the worthwhile difficulty their format cannot encompass. They sap the strength and vitality of a nation by rewarding conformity and quick-witted mediocrity while penalizing depth, subtlety, individuality, and creativity" (Smith, 1979).

A major problem with pre-college testing is that the scores are perceived by most students and the public as determining whether one student is smarter than the other (de Veer, 1981). The public and many admissions officers alike fail to recognize that the SAT and other tests are meant,at most, only to help predict the performance of the student in the first year of college, and the scores, when used alone are sometimes dead wrong as indicators. Unfortunately,a difference of only 10 to 20 points in the critical range (520 - low 600's ) may draw the line between attending a prestigious eastern Ivy League school and a small college in Pennsylvania (de Veer, 1981). In the lower range it might make the difference between attending or not attending any college.

A statistical study which I completed several years ago dealing with a two-year state university in Pennsylvania showed that with regard to the students who took the CEEB-SAT mathematics tests, high school mathematics average was a more valid predictor of college mathematics achievement than were the CEEB-SAT mathematics test scores. One reason for this,as suggested by the study, is the time factor involved in the standarized test. Many potentially good mathematics students would be able to work mathematical problems on a high school or

college mathematics test since the instructor generally gives sufficient time for the problems to be worked and checked. However, the nationally-normed tests might rush students who had the needed mathematical ability but could not perform successfully in an allotted time period. Also,as noted above, there is the question of the format of the testing itself. Lastly, and most importantly, the study considered performance in all math courses, including those beyond the freshman level.

The purpose of the present study, which is the focus of this paper, was to determine if the decline of the CEEB Scholastic Aptitude Test scores for mathematics is really a reflection of a decrease in our students' mathematical ability ,or the inadequacy of the test in measuring the student's mathematical ability. For this study random samples of 100 students were taken from the graduating classes of 1968, 1973, 1978, 1981, and 1982 from a liberal arts university in Northeastern Pennsylvania. In addition, the total 1983 graduating class of 508 was analyzed as well as the cumulative total of all classes mentioned. The data consisted of SAT mathematics test scores as well as the total mathematics grades of each student over his four-year college experience. The letter grades were converted to quality points for the respective three and four credit mathematics courses. A quality point as defined by this university is calculated by multiplying the Q.P.I. (A=4.0 scale) by the course credit. Thus an A for a 3 credit course would be 4.0 x 3 or 12 quality points whereas an A for a four credit course would equal 16 quality points and so on. Calculations were performed to determine if any relationships existed between CEEB SAT mathematics scores and

overall college mathematics achievement. Standard deviations and coefficients of determination were also calculated. The results of the various calculations are as follows; 1) the CEEB SAT mathematics scores have yielded extremely low correlations with each student's average math grade. This result may mean that high or low SAT math scores do not predict college mathematics performance, 2) extremely low correlations were prevalent over the fifteen year period under consideration (1968-1983). Those correlations were .026; .013; .116; .217; .153; and .089 for the years 1968, 1973, 1978, 1981, 1982, and 1983 respectively. 3) the total results for all years mentioned showed a very low correlation of .081 between CEEB SAT math scores and the average of four years of math grades. 4) For those students in the years considered, some correlations between CEEB SAT math scores and total four year math grades were found to be statistically significant at the .01 level of significance while others were not.

Some might contend that the reason for the low level of correlation is grade inflation and indeed grade inflation is a legitimate consideration. Between 1963 and 1974 the average undergraduate grade jumped from C+/B- to B and the percentage of "A" grades awarded more than doubled, from 16 percent to 34 percent (Suslow, 1977). However, since the present study indicates low correlations in either direction (high SAT scores/ low college performance and low SAT scores / high college performance) it seems that grade inflation is not a major factor. Ultimately, the study suggests quite strongly that the CEEB SAT math scores are not a good predictor of college performance.

TO CONCLUDE:

Unfortunately, what is all too often forgotten is that the SAT mathematics test score has its limitations, and is useful only to the extent that is is applied with understanding and discretion and with an intelligent appreciation of both its value and its limitations. The scores are, nevertheless, often used erroneously as exclusive measurements of individual or institutional quality. This, despite the fact that in 1975 CEEB President Sidney Marland stated that the tests were "not designed as a measure of school performance and should not be used that way." Two important principles concerning the SAT math test scores should not be overlooked, namely, 1) test scores, like all types of measurements, are not perfectly precise and should not be treated as though they were, 2) test scores are unreliable predictors of performance in college especially beyond the freshman level. Thus, students with high scores sometimes do worse in college than expected. Similarly, students with low scores sometimes do much better than expected.

There should be studies to develop and monitor a more valid procedure for understanding college achievement by the utilization of previous achievement measures for purposes of selection, placement, and guidance. Methods should be devised to measure personality traits and abilities not measured by nationally-normed tests, namely, outcomes of the educational process.

In addition, there should be a development of new test forms and procedures for educational measurements. Locating the

sources of error in widely used test forms would be another recommendation to improve the predictability factor of the nationally-normed tests.

Regarding the SAT itself, one should recall that the SAT was designed in 1941 as an "unchanging measurement" to predict student performance in college. Because of this the SAT might be considered a disguised achievement test which, in order to be valid, must better address what really matters in the two content fields which it measures.

REFERENCES


Anastasi, A. Psychological Testing, New York: Macmillan Publishing Comapny, 1976,p.171-172.

Biemiller, L. "SAT Math Scores Rise 3 Points in Math and One on Verbal Test," Chronicle of Higher Education (Sept. 26 1984):1+11.

Chissom, B.S. and Lanier, D. "Prediction of First Quarter Freshman GPA Using SAT Scores and High School Grades," Educational and Psychological Measurement (1975):461-63.

"College Board Frowns on SAT Comparisons," (UPI) The Scrantonian, (October 21, 1984).

Dalton, S. "A Decline in the Predictive Value of the SAT and High School Achievement," Educational and Psychological Measurement (1976):445-448.

de Veer, R.L. "Abuses and Uses of the SAT," Compact (Winter, 1981):8-11.

Fricke, B.J. College Entrance Examination Board Admission Testing Program. In O.K. Buros (Ed.), The sixth mental measurement yearbook. Highland Park, N.J.: Gryphon Press,1965. p.760.

Gussett, J.C. "College Entrance Examination Board Scholastic Aptitude Test Scores as a Predictor For College Freshmen Mathematics Grades," Educational and Psychological Measurement (1974),#34,953-955.

Larson, J.R. and Scontrino, M.P. "The Consistency of High School Grade Point Average and of the Verbal and Mathematical Portions of the Scholastic Aptitude Test of the College Entrance Examination Board as Predictors of College Performance: An Eight Year Study," Educational and Psychological Measurement (1976),#36,439-443.

Mauger, P. and Kolmodin, C.. "Long-term Predictive Validity of the Scholastic Aptitude Test," Journal of Educational Psychology (1975),#67,847-851.

Oldenquist, A. "The Decline of American Education in the 60's and 70's," American Education (May,1983):12-13.

Ryor, J. "Declining SAT Scores," Today's Education (1977):1-8.

Siegelman, M. "SAT and High School Average Predictions of Four Year College Achievement," Educational and Psychological Measurement (1971),#31,947-950.

Smith, R. J. "'Truth in Testing' Attracts Diverse Support," Science (1979),#205,1110-1114.

Suslow, S. "Grade Inflation: End of a Trend?" Change (March, 1977):17-18.

Thresher, B. A. Issues in School Measurement and Evaluation New York: Macmillan Publishing Co.1970,1-4.

Traxler, A. "Standarized Tests," National Educational Association Journal (November,1959):1-5.

Whitla, D. K. College Entrance Examination Board Admissions Testing Program. In O.K. Buros (Ed.), The sixth mental measurement yearbook. Highland Park, N.J.: Gryphon Press,1965. p.760.

Wirtz, W. "The Academic Score Decline: Are Facts the Enemy of Truth?" Phi Delta Kappan (1977):83-86.

Zimmerman, W.S.. College Entrance Examination Board Scholastic Aptitude Test. In O.K. Buros (Ed.), The sixth mental measurement yearbook. Highland Park, N.J.: Grypon Press, 1965. p.449.